MATH 217 W24 - LINEAR ALGEBRA, Section 001 (Dr. Paul Kessenich)
Homework Set Part B due ??? at 11:59pm
Zhengyu James Pan (jzpan@umich.edu)

1. Let W be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = (\vec{v}_1, ..., \vec{v}_d)$ be a basis for W. Consider the transformation $\mathbb{R}^n \xrightarrow{\pi} \mathbb{R}^n$ defined by

$$\pi(\vec{v}) = \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i} \vec{v}_i.$$

(a) Show that if $\vec{v}_i \cdot \vec{v}_j = 0$ for all $1 \leq i \neq j \leq d$, then the transformation π is the orthogonal projection onto W. (Note: this is almost, but not quite, the way we defined orthogonal projection. Make sure you understand how our definition is different from this before you start trying to prove it!)

Solution: Definition of Orthogonal Projection (from Definitions handout). If W is a subspace of an inner product space V and if $\vec{v} \in V$, the orthogonal projection of $\vec{v}$ onto W is the unique vector $\vec{w} \in W$ such that $\vec{v} - \vec{w} \in W^\perp$. The orthogonal projection of $\vec{v}$ onto W is sometimes denoted $\text{proj}_W(\vec{v})$.

Assume $\vec{v}_i \cdot \vec{v}_j = 0$ for all $1 \leq i \neq j \leq d$. Then

$$v - \pi(v) = v - \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i} \vec{v}_i.$$

So then for any $\vec{v}_n \in \mathcal{B}$,

$$\begin{aligned}
(\vec{v} - \pi(\vec{v})) \cdot \vec{v}_n &= \left(\vec{v} - \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i} \vec{v}_i \right) \cdot \vec{v}_n \\
&= \vec{v} \cdot \vec{v}_n - \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i} (\vec{v}_i \cdot \vec{v}_n) \\
&= \vec{v} \cdot \vec{v}_n - \sum_{i=1}^{d} \frac{\vec{v} \cdot \vec{v}_i}{\vec{v}_i \cdot \vec{v}_i} (\vec{v}_i \cdot \vec{v}_n)
\end{aligned}$$

Since we know $\vec{v}_i \cdot \vec{v}_j = 0$, this simplifies to

$$(\vec{v} - \pi(\vec{v})) \cdot \vec{v}_n = \vec{v} \cdot \vec{v}_n - \frac{\vec{v} \cdot \vec{v}_n}{\vec{v}_n \cdot \vec{v}_n} (\vec{v}_n \cdot \vec{v}_n) = \vec{v} \cdot \vec{v}_n - \vec{v} \cdot \vec{v}_n = 0.$$

An arbitrary vector $\vec{w}$ in W can be represented as a linear combination of the vectors in $\mathcal{B}$, $w = c_1\vec{v}_1 + ... + c_d\vec{v}_d$. So by distributivity of the dot product,

$$(\vec{v} - \pi(\vec{v})) \cdot w = c_1(\vec{v} - \pi(\vec{v})) \cdot \vec{v}_1 + ... + c_d(\vec{v} - \pi(\vec{v})) \cdot \vec{v}_d = 0.$$

Thus $\vec{v} - \pi(\vec{v}) \in W^\perp$ for all $\vec{v} \in \mathbb{R}^n$. So the transformation $\pi(\vec{v})$ always outputs the orthogonal projection of $\vec{v}$ on W.

(b) Give a counterexample to show that if the basis vectors in $\mathcal{B}$ are not perpendicular to each other, then the linear transformation π defined above π is not orthogonal projection onto W.

Solution: Let $n = 3$, $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ (not perpendicular), $W = \text{span}(\mathcal{B})$. Let $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Then since $\vec{v} \in W$, $\text{proj}_W(\vec{v}) = \vec{v}$. But

$$\begin{aligned}
\pi(\vec{v}) &= \frac{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\
&= \frac{1}{1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \neq \vec{v}.
\end{aligned}$$

So when $\mathcal{B}$ is not pairwise perpendicular, $\pi(\vec{v})$ is not necessarily the orthogonal projection.

2. Let $\mathcal{O}_n \subseteq \mathbb{R}^{n\times n}$ denote the set of orthogonal $n \times n$ matrices. Determine whether each of the following statements is True or False, and provide a short proof (or a counterexample) of your claim.

   (a) $\mathcal{O}_n$ is a subspace of $\mathbb{R}^{n\times n}$.

   Solution: False, $\mathcal{O}_n$ does not include the zero matrix, since the zero matrix is not orthonormal.

   (b) If $A, B \in \mathcal{O}_n$, then $AB \in \mathcal{O}_n$.

   Solution: True, by Problem 5 of WS 18.

   (c) If $A \in \mathcal{O}_n$, then $A^2 \in \mathcal{O}_n$.

   Solution: True, by part (b).

   (d) If $A^2 \in \mathcal{O}_n$, then $A \in \mathcal{O}_n$.

   Solution: False. Consider $A = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$.

   (e) If $A \in \mathcal{O}_n$ and A^2 is the identity matrix, then A is symmetric.

   Solution: True. $A = AI_n = AAA^\top = A^2A^\top = A^\top$.

3. (a) Suppose that $\mathcal{B} = (\vec{b}_1, ..., \vec{b}_r)$ is an orthonormal basis of a subspace V of $\mathbb{R}^n$. Prove that for all $\vec{v}, \vec{w} \in V$, $[\vec{v}]_\mathcal{B} \cdot [\vec{w}]_\mathcal{B} = \vec{v} \cdot \vec{w}$

   Solution: By the key theorem, the matrix S^{-1} which converts from $\mathcal{B}$-coordinates to standard coordinates is equal to $\begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_r \end{bmatrix}$. Since $\vec{b}_1, ..., \vec{b}_r$ are orthonormal, then S^{-1} is orthogonal (By WS18 Problem 4c). So $S = (S^{-1})^{-1}$, the change of basis matrix from standard to $\mathcal{B}$-coordinates, is also orthogonal, since it is the inverse (and transpose) of an orthogonal matrix. Thus the change of coordinates transformation T_S which S represents is also orthogonal (by WS18 Theorem B). Then by the definition of orthogonal, dot product is preserved. So $[\vec{v}]_\mathcal{B} \cdot [\vec{w}]_\mathcal{B} = T_S(\vec{v})T_S(\vec{w}) = \vec{v} \cdot \vec{w}$.

   (b) Prove that if $\mathcal{B} = (\vec{b}_1, ..., \vec{b}_r)$ and $\mathcal{C} = (\vec{c}_1, ..., \vec{c}_r)$ are two orthonormal bases of V, then $S_{\mathcal{B}\to\mathcal{C}}$ is an orthogonal $r \times r$ matrix.

   Solution: In the previous part, we showed that the change-of-coordinates matrices and transformations between an orthonormal basis and the standard basis are orthogonal. We also know that $S_{\mathcal{B}\to\mathcal{C}} = S_{\mathcal{B}\to\mathcal{S}}S_{\mathcal{S}\to\mathcal{C}}$, where $\mathcal{S}$ is the standard basis. The matrix product $S_{\mathcal{B}\to\mathcal{S}}S_{\mathcal{S}\to\mathcal{C}}$ represents the composition of two orthogonal change-of-coordinates transformations: from $\mathcal{B}$-coordinates to standard coordinates, then from standard coordinates to $\mathcal{C}$-coordinates. Since the composition of two orthogonal transformations is orthogonal, then $S_{\mathcal{B}\to\mathcal{C}}$ is orthogonal.

4. Let A be an $n \times m$ matrix. Prove or disprove each of the following statements:
   (a) $(\ker A)^{\perp} = \operatorname{im} A^{\top}$.

   Solution: True by WS19 Problem 4a.

   (b) $\mathrm{Rank}(A) = \mathrm{Rank}(A^{\top}A)$.

   Solution: By rank nullity on A, $n =$ Rank $A+$ nullity A. By rank nullity on $A^{\top}A$, $n =$ Rank $A^{\top}A+$ nullity $A^{\top}A$.
   For $\vec{x} \in \ker(A)$,
   $$A^{\top}A\vec{x} = A^{\top}\vec{0} = \vec{0}.$$
   So $\ker(A) \subseteq \ker(A^{\top}A)$.
   For $\vec{x} \in \ker(A^{\top}A)$:
   $$\begin{aligned} A^{\top}A\vec{x} &= 0 \\ x^{\top}A^{\top}A\vec{x} &= 0 \\ (A\vec{x})^{\top}(A\vec{x}) &= 0 \\ |Ax|^2 &= 0 \end{aligned}$$
   So $\ker(A^{\top}A) \subseteq \ker(A)$, and $\ker(A) = \ker(A^{\top}A)$. Then by the rank nullity equations, $\mathrm{Rank}(A) = \mathrm{Rank}(A^{\top}A)$.

   (c) $\mathrm{Rank}(A) = \mathrm{Rank}(A^{\top})$.

   Solution: True by WS19 Problem 4b.
   (below is reproving for practice, no need to grade this)
   dim(source) = rank + nullity
   For subspace W of V, dim(V) = dim(W) + dim(W$^{\perp}$)
   source of $A = R^m$
   source of $A^T = R^n$
   m = dim(im(A)) + dim(ker(A))
   $n = dim(im(A^T)) + dim(ker(A^T))$
   $n = dim(im(A^T)) + dim(im(A)^{\perp})$
   $n = dim(im(A^T)) + dim(targetofA) - dim(im(A))$
   $n = dim(im(A^T)) + n - dim(im(A))$
   $dim(im(A)) = dim(im(A^T))$
   Rank $A =$ Rank A^T

   (d) $\mathrm{Rank}(A^{\top}A) = \mathrm{Rank}\,(AA^{\top})$.

   Solution: By part (b), $\mathrm{Rank}(A^{\top}) = \mathrm{Rank}(AA^{\top})$ and $\mathrm{Rank}(A) = \mathrm{Rank}(A^{\top}A)$. By part (c), $\mathrm{Rank}(A) = \mathrm{Rank}(A^{\top})$. So $\mathrm{Rank}(A^{\top}A) = \mathrm{Rank}\,(AA^{\top})$ is true by transitivity.

   (e) $\ker A = \ker AA^{\top}$.

Solution: False. The dimensions of A and $AA^\top$ have different width, so their kernels will not even occupy the same space.

Definition. If A and B are two subsets of $\mathbb{R}^n$, then we say $A \perp B$ if for all $\vec{x} \in A$ and for all $\vec{y} \in B$, $\vec{x} \cdot \vec{y} = 0$. (Note that in this definition that A and B do not need to be subspaces, just subsets.)

Definition. A subset $A \subseteq \mathbb{R}^n$ is called pairwise orthogonal if any two elements $\vec{x}, \vec{y} \in A$ are orthogonal. Such a pairwise orthogonal subset $A \subseteq \mathbb{R}^n$ is called maximally pairwise orthogonal if it is not possible to enlarge set A to obtain a pairwise orthogonal subset $A' \subseteq \mathbb{R}^n$ that strictly contains A.

5. Let $n \in \mathbb{N}$. We consider the vector space $\mathbb{R}^n$.

(a) Prove that for all $X, Y \subseteq \mathbb{R}^n$, if $X \perp Y$ then $\text{Span}(X) \perp \text{Span}(Y)$.

Solution: Let $X = \{\vec{x}_1, \vec{x}_2, ...\}$ and $Y = \{\vec{y}_1, \vec{y}_2, ...\}$. Then let $x = c_1\vec{x}_1 + c_2\vec{x}_2 + ...$ and $y = d_1\vec{y}_1 + d_2\vec{y}_2 + ...$ be arbitrary vectors in X and Y respectively, where $c_i, d_i \in \mathbb{R}$. Then

$$\begin{aligned}
\vec{x} \cdot \vec{y} &= (c_1\vec{x}_1 + c_2\vec{x}_2 + ...) \cdot (d_1\vec{y}_1 + d_2\vec{y}_2 + ...) \\
&= (c_1\vec{x}_1) \cdot (d_1\vec{y}_1 + d_2\vec{y}_2 + ...) + (c_2\vec{x}_2) \cdot (d_1\vec{y}_1 + d_2\vec{y}_2 + ...) + ... && \text{(distribute)} \\
&= (c_1\vec{x}_1) \cdot (d_1\vec{y}_1) + (c_1\vec{x}_1) \cdot (d_2\vec{y}_2) + ... + (c_2\vec{x}_2) \cdot (d_1\vec{y}_1) + (c_2\vec{x}_2) \cdot (d_2\vec{y}_2) + ... \\
&= c_1d_1\vec{x}_1 \cdot \vec{y}_1 + c_1d_2\vec{x}_1 \cdot \vec{y}_2 + ... + c_2d_1\vec{x}_2 \cdot \vec{y}_1 + c_2d_1\vec{x}_2 \cdot \vec{y}_2 + ... \\
&= 0 + 0 + ... + 0 + 0 + ... && (x \cdot y = 0) \\
&= 0
\end{aligned}$$

So if $X \perp Y$, then also $\text{Span}(X) \perp \text{Span}(Y)$.

(b) Let X and Y each be a linearly independent subset of $\mathbb{R}^n$. Prove that if $X \perp Y$, then $X \cup Y$ is linearly independent.

Solution: Assume $X \perp Y$, and X and Y are both linearly independent. We know that no vector in X or Y is $\vec{0}$, since that would create a nontrivial relation. Then let a vector in $\text{Span}(X \cup Y)$ be $\vec{v} = c_1\vec{x}_1 + c_2\vec{x}_2 + ... + d_1\vec{y}_1 + d_2\vec{y}_2 + ... = \vec{0}$, where $c_i, d_i \in \mathbb{R}$. Note that $\vec{v}$ can also be expressed as the sum of a vector $\vec{x} = c_1\vec{x}_1 + c_2\vec{x}_2 + ...$ in $\text{Span}(X)$ and $\vec{y} = d_1\vec{y}_1 + d_2\vec{y}_2 + ...$ in $\text{Span}(Y)$. We want $\vec{v}$ to represent a relation on $X \cup Y$, so we set $\vec{v} = \vec{0}$. Then also $|\vec{v}|^2 = 0$. But

$$\begin{aligned}
\vec{v} \cdot \vec{v} &= (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) \\
&= |\vec{x}|^2 + 2(\vec{x} \cdot \vec{y}) + |\vec{y}|^2 \\
&= |\vec{x}|^2 + 2 \cdot 0 + |\vec{y}|^2
\end{aligned}$$

So a relation only exists when the components of v in $\text{Span}(X)$ and $\text{Span}(Y)$ are both $\vec{0}$. But this only occurs when all of c_i, d_i are 0, since X and Y are linearly independent. So $X \cup Y$ is linearly independent.

(c) Prove that every maximally pairwise orthogonal set of vectors in $\mathbb{R}^n$ has $n + 1$ elements.

Solution:

Lemma. The elements of a pairwise orthogonal set of vectors must be linearly independent with the exception of $\vec{0}$. That is, only the trivial relation can exist on the nonzero vectors of a pairwise orthogonal subset of vectors in $\mathbb{R}^n$.

Proof. Let X be an orthogonal set of vectors $\{x_1, x_2, ...\}$ in $\mathbb{R}^n$. Assume a nontrivial relation $c_1x_1 + c_2x_2 + ... = \vec{0}$ exists, where . Then $\vec{x}_i \cdot \vec{0} = \vec{x}_i \cdot (c_1x_1 + c_2x_2 + ...) = \vec{x}_i \cdot (c_ix_i) + 0 + 0 + ... = c_i|x_i|^2$ reveals a contradiction: $x_i = \vec{0}$. So only the trivial relation can exist on the nonzero vectors of a pairwise orthogonal subset of vectors in $\mathbb{R}^n$.

We already know that bases of $\mathbb{R}^n$ are the maximal linearly independent sets in $\mathbb{R}^n$, and have n elements. If a pairwise orthogonal set X does not include a basis, it is not maximal since there exist nonzero vectors in the orthogonal complement of $\text{Span}(X)$, which could be added to X and maintain its pairwise orthogonality. So any maximally pairwise orthogonal subset of $\mathbb{R}^n$ must include a basis of $\mathbb{R}^n$.

Since $\vec{0}$ has a dot product of 0 with any vector, it will be pairwise orthogonal with any set of vectors.

Thus any maximally pairwise orthogonal set of vectors in $\mathbb{R}^n$ is $\{\vec{0}\} \cup \mathcal{B}$, where $\mathcal{B}$ is an orthogonal basis of $\mathbb{R}^n$. No other vectors can be added, since they will be in the span of $\mathcal{B}$, and break the linear independence if added. Since $\mathcal{B}$ has n elements, a maximally pairwise orthogonal set of vectors in $\mathbb{R}^n$ will have $n + 1$ elements.

6. Let A be an $n \times m$ matrix, with $m \leq n$.

   (a) If $\text{rank}(A) = m$, prove that it is always possible to write $A = QL$, where Q is an $n \times m$ matrix with orthonormal columns and L is a lower triangular $m \times m$ matrix with positive diagonal entries.

   Solution: Let $\mathcal{U} = \{\vec{u}_1, ...\vec{u}_m\}$ be an orthonormal basis of $\text{im}(A)$ derived by descending Gram-Schmidt from the columns $\vec{a}_i$ of A (which we know are a basis $\mathcal{A}$ of $\text{im}(A)$ since $\text{rank}(A) = m$). That is, set $u_m = \frac{\vec{a}_m}{|\vec{a}_m|}$, and for each $1 \leq k < m$ let

$$\vec{u}_k = \frac{\vec{w}_k}{|\vec{w}_k|} \text{ where } \vec{w}_k = \vec{a}_k - \sum_{i=m}^{k+1} (\vec{a}_k \cdot \vec{u}_i)\vec{u}_i$$

   Note that this summation descends from a_m to a_1. Then let $Q = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}$. So Q has dimension $n \times m$ and has orthonormal columns.

   We know that for any two bases $\mathcal{A}$ and $\mathcal{B}$ for W, $B = AS_{\mathcal{B} \to \mathcal{A}}$, where A and B are matrices with the basis vectors as columns by WS17 Problem 6c. This applies to $\mathcal{A}$ and $\mathcal{U}$ as bases of $\text{im}(A)$: $A = QS_{\mathcal{A} \to \mathcal{U}}$. So, we just need to show that $S_{\mathcal{A} \to \mathcal{U}}$ is lower triangular.

$$S_{\mathcal{A} \to \mathcal{U}} = \begin{bmatrix} [a_1]_{\mathcal{U}} & \cdots & [a_m]_{\mathcal{U}} \end{bmatrix}$$

   By our definition of the Gram-Schmidt process to find $\mathcal{U}$,

$$\begin{aligned}
\vec{a}_k &= \vec{w}_k + \sum_{i=m}^{k+1} (\vec{a}_k \cdot \vec{u}_i)\vec{u}_i \\
\vec{u}_k \cdot \vec{a}_k &= \vec{u}_k \cdot (\vec{w}_k + \sum_{i=m}^{k+1} (\vec{a}_k \cdot \vec{u}_i)\vec{u}_i) \\
&= \frac{\vec{w}_k \cdot \vec{w}_k}{|w_k|} + \sum_{i=m}^{k+1} (\vec{a}_k \cdot \vec{u}_i)\vec{u}_i \cdot \vec{u}_k \\
&= |\vec{w}_k| + 0 \\
&= |\vec{w}_k|
\end{aligned}$$

   So then the kth column of $S_{\mathcal{A} \to \mathcal{U}}$ is

$$[a_k]_{\mathcal{U}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ a_k \cdot \vec{u}_k \\ a_k \cdot \vec{u}_{k+1} \\ \vdots \\ \vec{a}_k \cdot \vec{u}_{m-1} \\ \vec{a}_k \cdot \vec{u}_m \end{bmatrix}$$

   Since $\vec{u}_k \cdot \vec{a}_k = |\vec{w}_k|$, we know the diagonal elements of $L = S_{\mathcal{A} \to \mathcal{U}}$ are always positive, since $\vec{w}_k$ is nonzero. Additionally, every element above the diagonal of L is zero, so it is lower triangular. Thus we have found $A = QL$ which satisfy the conditions.

(b) Prove that if $\text{rank}(A) < m$, it is still possible to obtain such a decomposition if we allow some diagonal entries to be zero.

Solution: